

# **Document made available under the Patent Cooperation Treaty (PCT)**

International application number: PCT/IL05/000230

International filing date: 24 February 2005 (24.02.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US  
Number: 60/547,436  
Filing date: 26 February 2004 (26.02.2004)

Date of receipt at the International Bureau: 17 March 2005 (17.03.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland  
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

PCT/IL 2005 /000230

08 MAR 2005

PA 1281573

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

February 09, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM  
THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK  
OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT  
APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A  
FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: 60/547,436

FILING DATE: February 26, 2004

By Authority of the  
COMMISSIONER OF PATENTS AND TRADEMARKS



M. SIAS  
Certifying Officer





022604

PTO/SB/16 (08-03)  
Approved for use through 07/31/2006. OMB 0651-0032

U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.  
**PROVISIONAL APPLICATION FOR PATENT COVER SHEET**

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No.

INVENTOR(S)			
Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)	
NIR DAVID	CARMI GIDONI	GIVAT BRENER, ISRAEL MAZKERET BATIA, ISRAEL	
Additional inventors are being named on the _____ separately numbered sheets attached hereto			
TITLE OF THE INVENTION (500 characters max)			
COMPOSITIONS AND METHODS FOR ASYMMETRIC RECOMBINATION			
Direct all correspondence to: CORRESPONDENCE ADDRESS			
<input checked="" type="checkbox"/> Customer Number:	24505		
OR			
<input type="checkbox"/> Firm or Individual Name	ALPHAPATENT ASSOCIATES LTD.		
Address	P.O.B. 2345		
Address			
City	BEIT SHEMESH	State	Zip 99544
Country	ISRAEL	Telephone 516-620-4572 (US)	Fax 516-620-4572 (US)
ENCLOSED APPLICATION PARTS (check all that apply)			
<input checked="" type="checkbox"/> Specification Number of Pages 32	<input type="checkbox"/>	CD(s), Number _____	
<input checked="" type="checkbox"/> Drawing(s) Number of Sheets 7	<input type="checkbox"/>	Other (specify) _____	
<input type="checkbox"/> Application Date Sheet. See 37 CFR 1.76			
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT			
<input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27.	FILING FEE Amount (\$)		
<input type="checkbox"/> A check or money order is enclosed to cover the filing fees.			
<input checked="" type="checkbox"/> The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: 501380	\$160.00		
<input type="checkbox"/> Payment by credit card. Form PTO-2038 is attached.			
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.			
<input checked="" type="checkbox"/> No.			
<input type="checkbox"/> Yes, the name of the U.S. Government agency and the Government contract number are: _____			

[Page 1 of 2]

Respectfully submitted,

SIGNATURE \_\_\_\_\_

TYPED or PRINTED NAME DANIEL J. SWIRSKY

TELEPHONE 516-620-4572 (US)

Date February 26, 2004

REGISTRATION NO. 45,148

(if appropriate)

Docket Number: 1295-USP

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT  
 This collection of information is required by 37 CFR 1.51. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 8 hours to complete, including the gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden should be sent to the Chief Information Officer U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1-50. DO NOT SEND FEES OR COMPLETED -OIRMS TO THIS ADDRESS. SEND TO: Mail Stop Provisional Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

**PROVISIONAL APPLICATION COVER SHEET**  
**Additional Page**

PTO/SB/16 (08-03)

Approved for use through 07/31/2006. OMB 0651-0032  
U Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to collection of information unless it displays a valid OMB control number.

Docket Number                    1295-USP

**INVENTOR(S)/APPLICANT(S)**

Given Name (first and middle [if any])	Family or Surname	Residence (City and either State or Foreign Country)
YORAM PETER STEPHEN W.	EYAL SCHULTZ SANTORO	MESILAT ZION, ISRAEL SAN DIEGO, CALIFORNIA BOSTON, MASSACHUSETTS

[Page 2 of 2]

Number 2 of 2

**WARNING:** Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

## **COMPOSITIONS AND METHODS FOR ASYMMETRIC RECOMBINATION**

### **FIELD OF THE INVENTION**

The present invention relates to compositions and methods for catalyzing asymmetric recombination of non-palindromic recombination sites in a cell free system, in isolated cells or in living organisms. The compositions and methods of the invention are suitable for mediating specific recombinations between DNA sequences comprising recombination sites without being limited to strict palindromic symmetry within each recombination site.

### **10 BACKGROUND OF THE INVENTION**

Site-specific recombination systems mediate control of a large variety of critical biological functions in nature, through accurate excision, inversion or integration of defined DNA sequences. Site-specific recombination systems function through specific interactions of recombinase enzymes with their corresponding DNA target sequences, and provide an experimental model for studying protein-DNA interactions and protein-protein interactions among others.

Two of the most characterized recombinases are the Flp protein of yeast and the Cre protein of bacteriophage P1. These recombinases initiate recombination by binding to a specific recognition site, the *frt* site for Flp protein the *lox* site for the Cre protein. These two different recognition sites share a strict palindromic organization consisting of two identical 13-bp inverted repeats surrounding a spacer region, 8 bp in length, that confers directionality to the site and hence to the recombination reaction. Each of the palindrome halves provides a binding site for one recombinase monomer, while the recombination reaction takes place within the spacer region. A complete recombinase-mediated recombination takes place between two recombination sites, each bound by two recombinase monomers. The resulting four monomers are brought together by protein-protein interactions to form a homotetrameric assembly of identical recombinase molecules.

Lee et al. (*Gene*, 216:55-65, 1998) analyzed the role of each nucleotide sequence of the spacer region in the recombination process by constructing *loxP* spacer mutants with single-base or double-base substitutions (e.g. SEQ ID NOS:1-33; Table 2), which may facilitate recombination with diverse efficiencies.

Shaikh et al. (J. Mol. Biol. 302:24-48, 2000) discloses non-palindromic recombination sites comprising sequences derived from *lox* and *fri* sites that can be recognized and cleaved by chimeric Cre/Flp recombinases. The chimeric Cre/Flp recombinases can not complete recombination of the non-palindromic sites, apparently due to lack of ligation of the cleavage sites.

A recent publication by part of the inventors of the present invention discloses Cre variants, designated CM1 and CM2 (Table 1), which recognize wild type and/or variant *loxP* sites (SEQ ID NOS:35-37), and a method for identifying recombinase variants by a directed evolution strategy (Santoro et al., Proc. Natl. Acad. Sci. USA, 99: 4185-4190, 10 2002).

U.S. Patent No. 6,465,254 discloses mutant *loxP* sites and methods of using thereof. However, recombination is performed only between two identical mutant *loxP* recombination sites.

Methods for recombination in plants using a nucleotide sequence flanked between two non-identical recombination sites are disclosed in US Patent Nos. 6,573,425 and 6,664,108. The '425 patent relates to methods of integrating into plants a nucleotide sequence flanked between two non-identical recombination sites thus preventing or greatly suppressing excision of said nucleotide sequence, post-integration, in the presence of a recombinase. The '108 patent relates to Agrobacterium-mediated transfer of T-DNA to a plant cell, wherein the T-DNA contains a viral replicon flanked by target sites for a site-specific recombinase.

There is an unmet need for recombination systems that are not limited to wild type recognition sites and moreover that are not restricted to palindromic symmetry of recognition sites, thus enabling recombination of any desired recombination site.

## 25 SUMMARY OF THE INVENTION

The present invention relates to compositions and methods for mediating successful recombination events with increased site specificity. The compositions and methods of the present invention are directed for catalyzing asymmetric recombination at a predetermined genomic locus without being limited to a particular palindromic organization.

30 The present invention is based in part on the unexpected finding that a composition comprising two distinct recombinase proteins is capable of catalyzing recombination of non-

palindromic recombination sites with high efficiency. The recombination efficiency of such composition was found to be significantly higher than the recombination efficiency of a composition consisting exclusively or predominantly one of the recombinases.

The present invention provides means for recombination that are advantageous over previously known means for recombination. The main limitation of recombination systems known in the art is the requirement for precise palindromic recombination sites. The major advantage of recombination according to the principles of the present invention is that the recombination sites need not be palindromic sites. Rather, the recombination sites may be any desired sequence with the proviso that it can be recognized by at least one recombinase for the purpose of recombination.

Unexpectedly, asymmetric recombination between two recombination sites is carried out according to the present invention by a plurality of distinct recombinases. Without wishing to be bound by any particular theory or mechanism of action, the asymmetric recombination may be attributed to the ability of the plurality of different recombinases to form a heterotetramer thus, by bringing the two recombination sites together, facilitating recombination.

According to one aspect, the present invention provides a genetically modified host cell transformed by an isolated exogenous DNA molecule wherein the isolated exogenous DNA molecule is integrated by asymmetric recombination into a defined locus of the genome of the cell, said isolated exogenous DNA molecule comprising a polynucleotide sequence flanked by two recombination sites, at least one recombination site is an asymmetric recombination site.

According to yet another aspect, the present invention provides a genetically modified host cell, the cell lacking at a defined genomic locus an endogenous polynucleotide sequence wherein said cell comprising a first recombination site at the defined genomic locus, with the proviso that the endogenous polynucleotide sequence is flanked by the first and a second recombination sites, such that at least one of the recombination sites is an asymmetric recombination site.

The terms "asymmetric recombination site" or "chimeric recombination site" as used herein are interchangeable and used to describe a non-palindromic DNA element comprising a first and a second DNA sequences, also termed hereinafter non-palindromic halves. The two non-palindromic halves flank a spacer region which confers directionality

to the recombination site and hence to the recombination reaction. The first and second DNA sequences correspond to two recombinase recognition sites. In one embodiment, the two sites are recognized by at least one recombinase. In an alternative embodiment, the two sites are recognized by a plurality of recombinases. In yet another embodiment, at least one 5 non-palindromic half is not similar to a natural recognition site, such as the natural *frt* or *loxP* sites.

As used herein, the term "asymmetric recombination" refers to recombination between two recombination sites, wherein at least one of the recombination sites is an asymmetric recombination site. In one embodiment, the two recombination sites are asymmetric. In 10 another embodiment, the two recombination sites are the same asymmetric recombination sites. In yet another embodiment, the two recombination sites are distinct from each other.

It is to be understood explicitly that the scope of the present invention encompasses any form of recombination including, without limitation, asymmetric recombination between recombination sites that are in a *cis* or *trans* location. In the former situation, the 15 orientation of the recombination sites may be the same or the opposite. In the case of *trans* localization, the DNA strands involved can be linear or circular. The outcome of the asymmetric recombination according to the present invention may be excision or inversion of an intervening sequence in the case of *cis* location of two recombination sites and may be insertion of one DNA into another or translocation between two DNA molecules in the case 20 of *trans* located recombination sites.

The term "spacer" is to be construed in its most general sense and refers to an asymmetric core sequence consisting of 8 bp sequence located between the two half of a recombination site.

According to another embodiment, the present invention provides a genetically 25 modified host cell transformed by an isolated exogenous DNA molecule wherein the isolated exogenous DNA molecule is integrated by asymmetric recombination into a defined locus of the genome of the cell the isolated exogenous DNA molecule is specifically integrated into at least one locus of the genome of genetically modified host cell.

According to another embodiment, the isolated exogenous DNA molecule is 30 integrated into sites of the genome chosen from the group consisting of: 3' UTRs, 5' UTRs, polyA sites and gene promoters.

According to yet another embodiment, the host cell is selected from the group consisting of: yeast, plant cell, embryonic stem cell, mesenchymal cell, and haematopoietic progenitor cell.

According to yet another embodiment, the present invention provides a transgenic 5 organism, comprising said host cell, selected from the group consisting of: plant, yeast, or a vertebrate.

According to yet another embodiment, the host cell is further transformed by at least one expression vector comprising a polynucleotide sequence encoding at least one recombinase capable of catalyzing the asymmetric recombination.

10 According to yet another embodiment, the host cell is further transformed by at least one expression vector comprising a polynucleotide sequence encoding at least two distinct recombinases capable of catalyzing the asymmetric recombination, alternatively at least three distinct recombinases capable of catalyzing the asymmetric recombination, and alternatively at least four distinct recombinases capable of catalyzing the asymmetric 15 recombination.

According to yet another embodiment, the host cells is capable of expressing the at least one recombinase.

According to yet another embodiment, the at least one expression vector is a plasmid or a virus.

20 According to yet another embodiment, the polynucleotide sequence encoding the at least one recombinase is operably linked to expression control sequences.

According to yet another embodiment, the polynucleotide sequence encoding the at least one recombinase further comprising a promoter, the promoter is derived from bacteria, yeast, insect, animal, plant or virus. The promoter may be selected from the group 25 consisting of: *E. coli lac* and *trp* operons, the *tac* promoter, the bacteriophage $\lambda$ L promoter, bacteriophage T7 and SP6 promoters,  $\beta$ -actin promoter, insulin promoter, human cytomegalovirus (CMV) promoter, HIV-LTR, RSV-LTR, SV40 promoter, baculoviral polyhedrin and p10 promoter.

According to yet another embodiment, the promoter is an inducible promoter. The 30 inducible promoter may be selected from the group consisting of: tetracycline, heat shock,

steroid hormone, heavy metal, phorbol ester, adenovirus E1A element, interferon, and serum inducible promoters.

According to a preferred embodiment, the host cell is transformed by an expression vectors comprising a polynucleotide sequence encoding a plurality of recombinases capable of catalyzing the asymmetric recombination.

According to yet another preferred embodiment, the isolated transgenic host cells is capable of expressing the plurality of recombinases.

According to yet another embodiment, each recombinase recognizes at least one half of at least one of the two recombination sites.

According to yet another embodiment, at least one recombinase of the plurality of recombinases is a mutant recombinase.

According to one embodiment, the at least one recombinase is a Cre mutant wherein preferably the spacer sequence is consisting of any of the following sequences: SEQ ID NOS. 1-34.

According to yet another aspect, the present invention provides a method for catalyzing asymmetric recombination between recombination sites, wherein at least one recombination site is an asymmetric recombination site. Asymmetric recombination according to the present invention may mediate excision, integration or inversion and may be applied for a large number of applications in a cell free system, in vitro and in vivo.

According to one embodiment, the present invention provides a method for genetically modifying cells by catalyzing a site-specific insertion of a selected DNA segment into at least one defined locus of the genome of the cells, comprising:

(a) providing a population of cells having a first endogenous recombination site at a defined genomic locus;

(b) introducing into the population of cells an isolated DNA comprising a selected DNA segment and further comprising a second recombination site;

(c) introducing into the population of cells at least one vector comprising a polynucleotide sequence encoding at least one recombinase, the at least one recombinase being capable of catalyzing recombination

30

between the first endogenous recombination site and the second recombination site;

- (d) obtaining site-specific insertion of the selected DNA segment at the defined genomic locus; and

- 5 (e) selecting a subpopulation of cells or a cell progeny from said population of cells, the subpopulation of cells is devoid of said at least one recombinase;

wherein said first endogenous recombination site is an asymmetric recombination site.

- According to one embodiment, the first endogenous recombination site and the second recombination site are the same asymmetric recombination sites.

According to another embodiment, the method comprising introducing into the population of cells at least one vector comprising a polynucleotide sequence encoding at least two distinct recombinases capable of catalyzing the asymmetric recombination, alternatively at least three distinct recombinases capable of catalyzing the asymmetric recombination, and alternatively at least four distinct recombinases capable of catalyzing the asymmetric recombination.

According to yet another embodiment, the polynucleotide sequence encoding the at least one recombinase is operably linked to expression control sequences.

- According to yet another embodiment, the at least one vector further comprises a regulatory nucleotide sequence such that expression of said at least one recombinase is produced by activating the regulatory nucleotide sequence.

According to yet another embodiment, the selected DNA segment is a gene for a structural protein, an enzyme, or a regulatory molecule.

- According to yet another embodiment, the at least one expression vector is a plasmid or a virus.

According to yet another aspect, the present invention provides a method for genetically modifying cells by catalyzing a site-specific excision of an endogenous DNA segment from a defined locus of the genome of the cells, comprising:

- (a) providing a population of cells comprising an endogenous DNA segment within a defined locus of the genome thereof, the selected endogenous DNA segment is flanked by two recombination sites having the same orientation, wherein at least one recombination site is asymmetric;
  - (b) introducing into the population of cells a DNA molecule comprising a polynucleotide sequence encoding at least one recombinase, the at least one recombinase is capable of catalyzing recombination between the two recombination sites;
  - (c) obtaining site-specific excision of the endogenous DNA segment from the defined locus of the genome of said population of cells;  
and
  - (d) selecting a subpopulation of cells or a cell progeny from said population of cells, the subpopulation of cells is devoid of said at least one recombinase.

According to another embodiment, the DNA molecule further comprises a regulatory nucleotide sequence and expression of the at least one recombinase is produced by activating the regulatory nucleotide sequence.

According to yet another embodiment, the selected endogenous DNA segment is a gene for a structural protein, an enzyme, or a regulatory.

According to yet another embodiment, the two recombination sites have opposite orientations and the site-specific recombination is a site-specific inversion of the nucleotide sequence of the selected endogenous DNA segment.

According to yet another aspect, the present invention provides a method for catalyzing asymmetric recombination, comprising:

- (a) providing a first composition comprising at least one DNA molecule, the  
at least one DNA molecule comprising two recombination sites, wherein  
at least one recombination site is an asymmetric recombination site;

(b) providing a second composition comprising at least one recombinase, the  
at least one recombinase is capable of catalyzing recombination of the  
two recombination sites; and

(c) contacting the first composition and the second composition to obtain asymmetric recombination between the two recombination sites.

According to another embodiment, the first composition comprises two DNA molecules, each DNA molecule comprising one recombination site wherein at least one DNA molecule comprises an asymmetric recombination site. In an alternative embodiment 5 at least one DNA molecule is circular.

According to another embodiment, the second composition comprises at least two distinct recombinases capable of catalyzing the asymmetric recombination, alternatively at least three distinct recombinases capable of catalyzing the asymmetric recombination, and 10 alternatively at least four distinct recombinases capable of catalyzing the asymmetric recombination.

According to yet another embodiment, each distinct recombinase specifically recognizes at least one half of a recombination site.

According to yet another embodiment, the first composition comprising a DNA 15 molecule comprising two recombination sites wherein at least one recombination site comprises SEQ ID NO:37 and the second composition comprising a CM2 Cre mutant (Table 1).

According to yet another aspect, the present invention provides methods of cell therapy using the disclosed genetically modified cells of the invention.

20 Other objects, features and advantages of the present invention will become clear from the following description and drawings.

#### BRIEF DESCRIPTION OF THE DRAWINGS

**Figure 1** demonstrates SEQ ID NOS.:35-37.

25 **Figure 2** presents a membrane (A) and a quantitative representation thereof (B) of in vitro recombination as a function of recombinase content: a. wt Cre, b. CM2, c. a mixture of 50% CM2 and 50% wt Cre, and d. CM1.

**Figure 3** shows in vitro recombination activity as a function of reaction time with three different total recombinase concentrations: A, 30 nM; B, 60 nM and C, 90 nM.

30 **Figure 4** exhibits SEQ ID NOS: 35 and 38.

**Figure 5** shows schematic representations (A-C) of a strategy for selecting Cre mutants capable of catalyzing asymmetric recombinations.

## **DETAILED DESCRIPTION OF THE INVENTION**

### **5    Definitions**

The terms "sequence-specific recombinase" and "site-specific recombinase" refer to enzymes that recognize and bind to a specific recombination site or sequence and catalyze the recombination of nucleic acid in relation to these sites.

The terms "sequence-specific recombinase target site" and "site-specific recombinase target site" refer to short nucleic acid site or sequence which is recognized by a sequence- or site-specific recombinase and which become the crossover regions during the site-specific recombination event. Examples of sequence-specific recombinase target sites include, but are not limited to, *lox* sites, *frt* sites, ATT sites and DIF sites. According to a currently preferred embodiment, the target sites are asymmetric recombination sites, wherein each asymmetric recombination site comprises a first and second non-palindromic halves flanking a spacer region which confers directionality to the recombination site and hence to the recombination reaction.

The term "*lox* site" as used herein refers to a nucleotide sequence at which the product of the *cre* gene of bacteriophage P1, Cre recombinase, can catalyze a site-specific recombination. A variety of *lox* sites are known to the art including the naturally occurring *loxP* (the sequence found in the P1 genome), *loxB*, *loxL* and *loxR* (these are found in the E. coli chromosome) as well as a number of mutant or variant *lox* sites such as *loxP511*, *loxΔ86*, *loxΔ117*, *loxC2*, *loxP2*, *loxP3* and *loxP23*. The *loxP* site comprises two 13 bp inverted repeat sequences separated by an 8 bp spacer region (Hoess et al., Proc. Natl. Acad. Sci. USA 79:3398, 1982). The internal spacer sequence of the *loxP* site is asymmetrical and thus, two *loxP* sites can exhibit directionality relative to one another (Hoess et al. Proc. Natl. Acad. Sci. USA 81:1026, 1984). When two *loxP* sites on the same DNA molecule are in a directly repeated orientation, Cre excises the DNA between these two sites leaving a single *loxP* site on the DNA molecule. (Abremski et al. Cell 32:1301, 1983). If two *loxP* sites are in opposite orientation on a single DNA molecule, Cre inverts the DNA sequence between these two sites rather than removing the sequence. The Cre recombinase also recognizes a

number of variant or mutant lox sites relative to the *loxP* sequence. Examples of these Cre recombination sites include, but are not limited to, the *loxB*, *loxL* and *loxR* sites which are found in the *E. coli* chromosome.

The term "frt site" as used herein refers to a nucleotide sequence at which the product 5 of the FLP gene of the yeast 2 micron plasmid, FLP recombinase, can catalyze a site-specific recombination.

As used herein, the term "vector" is used in reference to nucleic acid molecules that transfer DNA segment(s) from one cell to another. The term "vehicle" is sometimes used interchangeably with "vector." A "vector" is a type of "nucleic acid construct." The term 10 "nucleic acid construct" includes circular nucleic acid constructs such as plasmid constructs, plasmid constructs, cosmid vectors, etc. as well as linear nucleic acid constructs (e.g., λ-phage constructs, PCR products). The nucleic acid construct may comprise expression signals such as a promoter and/or an enhancer (in such a case it is referred to as an expression vector).

The term "expression vector" as used herein refers to a recombinant DNA molecule 15 containing a desired coding sequence and appropriate nucleic acid sequences necessary for the expression of the "operably linked" coding sequence in a particular host. Nucleic acid sequences necessary for expression in prokaryotes usually include a promoter, an operator (optional), and a ribosome binding site, often along with other sequences. Eukaryotic cells 20 are known to utilize promoters, enhancers, and termination and polyadenylation signals.

The scope of the present invention encompasses asymmetric recombination between 25 recombination sites that are in a *cis* or *trans* location. As used here, the term "*cis*" refers to genetic changes that are on the same DNA molecule in simple organisms or in the same haploid genome in cases where there are multiple chromosomes. The term "*trans*" refers to genetic changes that are carried by different genomes that have been introduced into the same cell (in any of the possible ways discussed above).

#### Preferred modes for carrying out the invention

The ability to direct site-specific recombination into natural sites in the eukaryotic 30 genome, using the existing means and methodologies, is limited particularly due to the strict prerequisite for palindromic recombination sites. The present invention provides

compositions and methods that provide a versatile solution to such limitation as recombination according to the present invention is enabled with the involvement of non-palindromic recombination sites.

Thus, the teaching of the present invention broadens the prospects for genetic manipulation of the eukaryotic genome, enabling integration, deletion or replacement of specific genes and DNA segments in defined genomic loci. Site directed recombination with at least one asymmetric recombination site, particularly endogenous asymmetric recombination sites, requires selection of recombinases that recognize the asymmetric recombination sites, and further capable of catalyzing the asymmetric recombination.

Asymmetric recombination of endogenous sites using a combination of recombinases may be utilized for treating a variety of diseases. For example, the working of the present invention may be utilized for excision of HIV provirus from the genome of infected CD4+ cells. Another therapeutic approach includes transforming human bone marrow cells with a combination of recombinases expression constructs. The transformed cells may constitutively or transiently produce the combination of active recombinases thus protecting the cells from HIV infection by an efficient excision of the asymmetric-flanked virus sequence upon infection.

Site-specific asymmetric recombination according to the present invention may be also utilized for genetically modifying plants as plants lack efficient homology recombination. Agrobacterium vector is commonly used for plant transformation, however due to random integration of the transgene the majority of transgenic plants confer low transgenic expression and, therefore, non-desired phenotype. Utilization of recombination according to the teaching of the present invention provides an improved methodology resulting in the efficient generation of desirable phenotypes.

It is to be understood that the site for insertion is selected prior to insertion and is used for screening for recombinase(s) that recognize said site at adequate specificity. The site for insertion may be identified in silico and must comprise a core spacer sequence that is similar to a recombinase core sequence. For example, screening for a Cre mutant that can catalyze the asymmetric recombination is performed using, an asymmetric recombination site comprising a spacer sequence having 70% homology, preferably 80% homology to a sequence selected from the group consisting of: SEQ ID NOS:1-34.

The present invention provides a method for catalyzing asymmetric recombination between recombination sites, wherein at least one recombination site is an asymmetric recombination site, comprising providing a population of cells, genetically modifying the cell by inserting into, or excising from, a defined locus of said cells a DNA segment. The 5 method further comprises selecting a subpopulation of the genetically modified cells which is devoid of the at least recombinase that is capable of performing said asymmetric recombination.

Selecting the desired subpopulation may be achieved by using a selectable marker gene carried by the recombinant expression vector or attached to the exogenous DNA which is 10 integrated into the genomic locus of the genetically modified cells. The expression vector may contain more than one selectable marker to facilitate subsequent identification and selection of clones of cells comprising the recombination product under suitable conditions. The selectable marker may encode any functional element, such as protein, peptide, RNA, binding site for RNA and proteins, or products that provide resistance to organic or inorganic 15 agents. Examples of selectable markers include, but are not limited to, reporter genes such as P-galactosidase (GAL), fluorescent proteins (e.g., GFP, GFP-UV, EFP, BFP, EBFP, ECFP, β-EYFP), secreted form of human placental alkaline phosphatase such as SEAP, β-glucuronidase (GUS); resistance genes that encodes products which provide resistance against other wise toxic agents such as antibiotics (e.g. neomycin G418, hygromycin 20 resistant gene and puromycin resistant gene), yeast selectable markers leu2-d and URA3, apoptosis resistant genes (e.g. the baculoviral p35 gene) that encode proteins that binds to products which are detrimental to cell survival and promote apoptosis; antisense-oligonucleotides, and any other DNA that encodes product that directly or indirectly confer sensitivity of cells to particular agents.

25 The recombinant expression vector or the exogenous DNA may optionally include an affinity tag for selection and isolation of protein product encoded by same. Examples of such an affinity tag include, but are not limited to, a polyhistidine tract, polyarginine, glutathione-S-transferase (GST), maltose binding protein (MBP), a portion of staphylococcal protein A (SPA), and various immunoaffinity tags (e.g. protein A) and epitope tags such as those 30 recognized by the EE (Glu-Glu) antipeptide antibodies. The affinity tag may also be a signal peptide either native or heterologous to baculovirus, such as honeybee mellitin signal peptide. The affinity tag may be positioned at either the amino- or carboxy-terminus of the donor DNA.

The promoter controls expression of the recombinase of interest and the selectable marker gene within the recombinant vector are introduced into a host cell. The vector may further contain a polyadenylation signal for expression in mammalian cells.

In the event that the method of the invention is for inserting by asymmetric recombination an exogenous DNA into a predetermined genomic location, the exogenous DNA sequence may be any deoxyribonucleotide sequence encoding a functional gene or any synthetically generated DNA sequence. For example the exogenous DNA segment may be a sequence derived from cDNA of a particular gene or one of the members of a cDNA library. The cDNA library may be produced by converting mRNAs in a sample into double-stranded complementary DNA (cDNA) by using reverse transcriptase (RT) and the Klenow fragment of nucleic acid polymerase I. Depending on the source of mRNA sample, the cDNA library may contain various populations of genes of interest, such as disease genes located in certain tissue or type of cells. The exogenous DNA may also be a genomic DNA that contains the coding region interrupted with non-coding sequences (introns/intervening sequences). These introns may contain regulatory elements such as enhancers

The exogenous DNA may further comprise a promoter sequence that controls the expression thereof. The promoter may be any array of DNA sequences that interact specifically with cellular transcription factors to regulate transcription of the downstream gene. The promoter may be derived from any organism, such as bacteria, yeast, insect and mammalian cells and viruses. The selection of a particular promoter depends on what cell type is to be used to express the protein of interest. Examples of the promoter include, but are not limited to, E. coli lac and trp operons, the tac promoter, the bacteriophage  $\lambda\beta^P L$  promoter, bacteriophage T7 and SP6 promoters,  $\beta$ -actin promoter, insulin promoter, human cytomegalovirus (CMV) promoter, HIV-LTR (HIV-long terminal repeat), Rous sarcoma virus RSV-LTR, simian virus SV40 promoter, baculoviral polyhedrin and p10 promoter. The promoter may also be an inducible promoter that regulates the expression of downstream gene in a controlled manner, such as under a specific condition of the cell culture. Examples of inducible promoters include, but are not limited to, the bacterial dual promoter (activator/repressor expression system) which regulates gene expression in mammalian cells under the control of tetracyclines (Gossen et al., Proc. Natl. Acad. Sci. USA, 89, 5547-5551, 1992) and promoters that regulate gene expression under the control of factors such as heat shocks, steroid hormones, heavy metals, phorbol ester, the adenovirus E1A element, interferon, or serum.

According to one embodiment of the present invention, the site-specific asymmetric recombination is performed by a plurality of recombinases which may comprise at least one distinct wild type recombinase. The wild type recombinase may be derived from prokaryotic and eukaryotic sources.

5 Examples of site-specific recombination include, the following: 1) chromosomal rearrangements which occur in *Salmonella typhimurium* during phase variation, inversion of the FLP sequence during the replication of the yeast 2 micron circle and in the rearrangement of immunoglobulin and T cell receptor genes in vertebrates; 2) integration of bacteriophages into the chromosome of prokaryotic host cells to form a lysogen; and 3)  
10 transposition of mobile genetic elements (e.g., transposons) in both prokaryotes and eukaryotes.

One half of the asymmetric recombination site of the invention may also be selected from a variety of other recombination sites recognized by recombinases other than Cre. Examples of the non-Cre recombinases include, but are not limited to, site-specific recombinases include: the Int recombinase of bacteriophage, the FLP recombinase of the 2pi plasmid of *Saccharomyces cerevisiae*, the resolvase family, transposase of *Bacillus thuringiensis*.

The Int recombinase of bacteriophage  $\lambda$  belongs to the integrase family and mediates the integration of the  $\lambda$  genome into the *E. coli* chromosome. The Int recombinase of bacteriophage  $\lambda$  promotes irreversible recombination between its substrate ATT sites as part of the formation or induction of a lysogenic state (Landy, A., Ann. Rev. Biochem. 58:913, 1989). Reversibility of the recombination reactions results from two independent pathways for integrative and excessive recombination. Each pathway uses a unique but overlapping set of the 15 protein binding sites that comprise ATT site DNAs. Cooperative and competitive interactions involving four proteins (Int, Xis, IHF and FIS) determine the direction of recombination. Integrative recombination involves the Int and IHF proteins and sites ATT-P (240 bp) and ATT-B (25 bp). Recombination results in the formation of two new sites: ATT-L and ATT-R. Excessive recombination requires Int, IHF, and Xis, and sites ATT-L and ATT-R to generate ATT-P and ATT-B. Derivatives of the ATT site with changes within the  
25 30 15 bp core may also be suitable for efficient recombination. Integrase can be obtained as described by Nash, H. A., (1983) Methods of Enzymology 100:210-216.

The other members of the Integrase family of site-specific recombinases may also be used to construct libraries of alternative recombination proteins and recombination sites for the present invention. Examples of such Int recombinases include, but not limited to, site-specific recombinase encoded by bacteriophage  $\lambda$ , phi 80, P22, P2, 186, P4. This group of recombinases exhibits a large diversity of sequences, but all of the recombinases can be aligned in their C-terminal halves. Three positions are perfectly conserved within this family: histidine, arginine and tyrosine are found at respective alignment positions 396, 399 and 433 within the well-conserved C-terminal region. These residues contribute to the active site of this family of recombinases, and suggest that tyrosine-433 forms a transient covalent linkage to DNA during strand cleavage and rejoining.

The FLP recombinase of the 2pi plasmid of *Saccharomyces cerevisiae* recognizes the *frt* site which, like the *loxP* site, comprises two 13 bp inverted repeats separated by an 8 bp spacer. The FLP gene has been cloned and expressed in *E. coli* and in mammalian cells and has been purified (e.g. Meyer-Lean et al. Nucleic Acids Res. 15:6469, 1987).

The resolvase family members, such as the Tn3 resolvase, the Hin recombinase, and the Cin recombinase, may also be used for recombination according to the present invention. Transposase of *Bacillus thuringiensis* may also be used as recombination proteins and recombination sites. *Bacillus thuringiensis* is an entomopathogenic bacterium whose toxicity is due to the presence in the sporangia of A-endotoxin crystals active against agricultural pests and vectors of human and animal diseases. Most of the genes coding for these toxin proteins are plasmid-borne and are generally structurally associated with insertion sequences.

Other recombination systems may also be used as recombination proteins and recombination sites, including the *xerC* and *xerD* recombinases of *E. coli* which together form a recombinase (e.g. Leslie et al., EMBO J. 14:1561, 1995).

Additional strategy for selection of the desired subpopulation of cells includes applying the cis-trans test. This test provides an assessment as to whether two mutations that occur in different genomes have altered the same unit of function (usually the coding sequence for a single protein). It is expected that when the two mutations are put together in the same genome (*cis*) that they will not support normal function, even if they occur in two separate functional units. Such a test is straightforward in haploid organisms. However, to see *cis* lack of function in a diploid organism, it is necessary to make the organism

homozygous for the *cis* genome. Otherwise there will be a wild type genome opposite it that will support normal function. In most cases, the *cis* part of the *cis-trans* is an implied control that could be run, rather than an important part of the actual experimental study.

According to yet another aspect, the present invention provides methods of cell therapy using the disclosed genetically modified cells of the invention. Traditional cell therapy approaches utilize ex vivo gene transfer, which involves the initial step of obtaining cells from a subject in need thereof, following transforming of cells in vitro with a desired DNA sequence and finally introducing the transformed cells back into the subject.

## 10 EXAMPLES

### Materials and Methods

Construction of wt and mutant *loxP* substrates: *lox* substrates were constructed as follows: a chimeric asymmetric substrate, *loxP-M7*, containing one half-site identical to *loxP* and a second half-site identical to *lox-M7* (Fig. 1A) was designed by annealing two oligonucleotides that contained two restriction sites, *Xhol/BamHI* or *EcoRI/PstI*, and ligating them into a 1-kb spacer region at the *BamHI/EcoRI* restriction sites (SEQ ID NOS: 39-41; Fig. 1B). The resulting constructs were cloned into a Bluescript<sup>TM</sup> vector at the *Xhol/PstI* restriction sites, in direct orientation.

Expression and purification of recombinant Cre versions: C2(+)#1, wt Cre, and C2(+-) #4 His-tag cloned plasmids (supplied by Santoro et al., *ibid*) were transfected into DH10B *E. coli* cells. 500 ml bacterial culture containing LB with ampicillin (100 mg/L) were grown to an OD<sub>600</sub> of 0.8-1, induced with 0.02% L-arabinose and incubated at 37°C for 3 h. The cells were lysed within a lysis buffer (NaH<sub>2</sub>PO<sub>4</sub> 50 mM, pH 8, NaCl 300 mM and imidazole 10 mM). Protein was bound to ProBond<sup>TM</sup> beads (Invitrogen), washed three times with wash buffer (NaH<sub>2</sub>PO<sub>4</sub> 50 mM, pH 8, NaCl 300 mM and imidazole 20 mM), eluted with elution buffer (NaH<sub>2</sub>PO<sub>4</sub> 50 mM, pH 8, NaCl 300 mM and imidazole 250 mM) on Poly-Prep Chromatography columns (Bio-Rad) and then dialyzed twice in 50% glycerol, 300 mM NaCl, 10 mM Tris (pH 7.5) and 0.1% 2-mercaptoethanol. Protein concentrations were determined by Bradford assay (Bio-Rad) and confirmed by analysis on a 10% SDS-PAGE gel.

Example 1 - *In vitro* analysis of Cre recombinases

Concentration assays were performed at concentrations of 15-180 nM for 1 h. The reaction buffer contained 300 mM NaCl, 20 mM Tris (pH 7.5) and 1 mM EDTA in a final volume of 40 µL. All reactions were incubated at 37°C and stopped by incubation at 70°C for 10 min. Time-course experiments were carried out for 15-90 min, with 60 or 30 nM of enzyme, a mixture of 30 nM of Cre-wt plus 30 nM of C2(+/-)#4, or 15 nM of Cre-wt plus 15 nM of C2(+/-)#4. Reaction products were analyzed on 1% agarose gel containing 0.3 µg/µL ethidium bromide and transferred to a Nytransupercharge nylon membrane. The plasmid containing *LoxP-M7* substrate was linearized with *NcoI* (Fig. 2A).

A sample of 50-ng (1.25 nM) of 4kb linear fragment containing *LoxP-M7* substrate was random-prime labeled with  $\alpha$ -<sup>32</sup>P-dATP and hybridized to a membrane overnight. A quantitative representation of the recombination was performed by exposing the sample to a phosphoimager (FLA-5000, Fuji) and analyzing with Image Gauge software V 4.0 (Fuji).

Figure 1A presents lox symmetric and asymmetric sites. The *loxP* wild-type (wt) site for Cre recombinase is a symmetric site (FIG. 1A; SEQ ID NO:35), and *lox M7* is a symmetric mutant version of *loxP* (FIG. 1A; SEQ ID NO:36). *loxP-M7* is a chimeric asymmetrical version comprising one 13-bp repeat from *loxP* and one from *lox M7* (FIG. 1A; SEQ ID NO:37). Nucleotides that differ between the versions are underlined, denoting nucleotides of the wild-type version (*loxP*) or are given in bold, denotes nucleotides of the *lox M7* mutant version. The spacer region is in gray.

Two Cre variants, CM1 and CM2, corresponding to C2(+/#1 and C2(+/-/#4 (Santoro et al., *ibid*), were selected for catalyzing the recombination of *lox M7*. CM1 and CM2 have five substituted amino acids compared with the wt Cre, whereas two of the five are identical in CM1 and CM2 (Table 1). CM1 was found to be less specific, and it catalyzed the recombination of *lox M7* and of *loxP* with similar specificity and efficiency, whereas CM2 was found to be much more specific to *lox M7* than to *loxP*.

Table 1. Amino acid residues that were modified in Cre mutants with respect to wt Cre:

Designation	Residues in wt Cre vs. residues in Cre mutants				
wt Cre	I174	T258	R259	E262	E266
CM1 (C2(+/#1)	L174	N258	S259	G262	G266
CM2 (C2(+/-/#4)	A174	L258	S259	H262	G266

Selected mutants of the Cre recombinases and of wt Cre were analyzed for their capability to facilitate recombination, as homotetrameric versus heterotetrameric structures on chimeric asymmetric site. For this purpose, two chimeric asymmetric *lox* site of *loxP* and *lox M7*, also called *loxP-M7* (Fig. 1A; SEQ ID NO:37), were cloned into a Bluescript<sup>TM</sup> vector at directed orientation, flanking a 1-kb DNA fragment (Fig. 2A). *loxP-M7* substrate (1.25 nM) cloned into Bluescript<sup>TM</sup> vector in directed orientation was linearized and incubated with Cre recombinase versions for 1 h at concentrations of 30, 60 and 90 nM. Samples were separated by electrophoresis on a 1% agarose gel and then subjected to Southern hybridization, with the entire linearized Bluescript<sup>TM</sup> *loxP-M7* plasmid (4 kb) used as a probe. The resulting membrane was exposed to film for 24 h. Recombinases analyzed for activity were a. wt Cre, b. CM2, c. a mixture of 50% CM2 and 50% wt Cre, d. CM1 (FIG. 2A). Arrows are pointing to the linearized plasmid containing the substrate (4 kb); the linear plasmid recombination product (3 kb); and the circular insert recombination product (~ 1 kb).

A quantitative representation of recombination activity from the membrane described in FIG. 2A was collected by a phosphoimager, and the band intensities were quantified. The results of recombination activity were calculated as combined intensities of 1 kb + 3 kb products divided by the intensity of the total substrate in the reaction (1 kb + 3 kb + 4 kb; FIG. 2B).

In vitro recombination activity as a function of reaction time with various concentrations and recombinases was also tested (FIGS. 3A-C). A sample of linear *loxP-M7* substrate, 1.25 nM, was incubated with Cre recombinase versions for 15, 30 or 60 min. Samples were separated by electrophoresis on 1% agarose gel and then subjected to Southern hybridization with the entire linearized Bluescript<sup>TM</sup> *loxP-M7* plasmid (4 kb) used as a probe. Data from the membrane were collected by phosphoimager, and band intensities were quantified. The results of recombination activity were calculated as combined intensities of 1 kb + 3 kb products divided by the intensity of the total substrate in the reaction (1 kb + 3 kb + 4 kb). Total recombinase(s) concentrations were: 30 nM (FIG. 3A), 60 nM (FIG. 3B) and 90 nM (FIG. 3C), wherein these concentrations refer to the concentration of each recombinase (wt Cre, CM1 and CM2) when provided alone. When a combination of recombinases was provided, then the combination contained equal amounts of each recombinase, namely, 15 nM CM2 and 15 nM wt Cre (FIG. 3A), 30 nM CM2 and 30 nM wt Cre (FIG. 3B) and 45 nM CM2 and 45 nM wt Cre (FIG. 3C).

The wt Cre did not exhibit any recombination activity on *loxP-M7* (FIGS. 2-3), whereas CM1 facilitated recombination on *loxP-M7* at the same level of efficiency as on *lox M7*. CM2 did facilitate recombination on *loxP-M7* as efficiently as CM1, but only at the higher concentration of 90 nM (FIG. 2). However, at 30 nM CM2 was five fold slower than  
5 a mixture of 15 nM CM2 with 15 nM wt Cre, or of 30 nM CM1 (FIG. 2). At an enzyme concentration of 60 nM, CM2 was 2-fold slower than a mixture of 30 nM CM2 and 30 nM wt Cre. CM1 or CM2 at enzyme concentrations of 90 nM, and a mixture of 45 nM CM2 and 45 nM wt Cre catalyzed recombination equally efficiently (FIGS. 2-3). The results indicate  
10 that the mixture of two Cre variants, CM2 and wt Cre, formed a heterotetramer structure with higher specificity to *loxP-M7* than either of the Cre variants alone, and therefore catalyzed recombination as efficiently as CM1 (FIGS. 2-3).

Crystal structure analysis revealed that a complex of protein-DNA with water molecules affected the binding specificity of CM2 to *lox M7*, and that the flexibility of the protein-DNA contacts of CM1 resulted from the substitution to Asn258 and Ser259, thus allowing less specificity, with consequently equally effective recognition by *lox M7* and *loxP* (Enoch et al., Chem. Biol. 10:1085, 2003). The present invention demonstrates for the first time that Cre variant catalyzes asymmetric *lox* site recombination and moreover a mixture of two Cre variants could catalyze asymmetric *lox* site recombination by forming a heterotetramer structure. At the rather higher concentration of 90 nM (FIG. 3C), CM2 efficiently catalyzed the recombination of *loxP-M7* (FIGS. 2-3). The *lox M7* half of the *loxP-M7* exhibited optimal affinity to CM2 and it is suggested that at the higher concentration of 90 nM one monomer of CM2 stabilizes the other monomer on the *loxP* half of the *loxP-M7*. Previous analysis of the crystallographic structure revealed that Cre acts on its *loxP* substrate by targeting of *loxP* site with two Cre monomers (Guo et al., Nature, 389: 40-46, 1997) therefore it is suggested that four Cre monomers can catalyze the recombination of two symmetric *loxP* sites via an intermediate synapse of protein-protein interaction as a tetramer structure. Mutation within the inverted repeat sequence of residues at positions 1 and 8-13 that disrupts the palindromic structure appeared to have some tolerance for Cre recombinase, but residues at position 2-6 exhibited a very low degree of flexibility.  
20  
25

The spacer sequence has the cleavage-ligation site on both strands. Although this sequence is restricted, some mutated spacers could be recognized by wt Cre (Lee et al., Gene 216:55-65, 1998 and Table 2).

Table 2 – Mutants of the *lox* spacer (Lee et al., Gene, *ibid*) and wild-type *lox* spacer

Sequence	SEQ ID NO
ACGTATGC	1
AAGTATGC	2
AGGTATGC	3
ATATATGC	4
ATCTATGC	5
ATTATATGC	6
ATGCATGC	7
ATGAATGC	8
ATGGATGC	9
ATGTGTGC	10
ATGTTTGC	11
ATGTCTGC	12
ATGTACGC	13
ATGTAGGC	14
ATGTATAAC	15
ACGTATGC	16
AAGTATGC	17
AGGTATGC	18
ATGTATAAC	19
ATGTATCC	20
ATGTATTTC	21
ACGTATAAC	22
ATATATAAC	23
ATGCATAAC	24
ATGTGTAC	25
AAGTATCC	26
ATCTATCC	27
ATGAATCT	28
ATGTTTCC	29
AGTTATTTC	30
ATTTATAAC	31
ATGGATTTC	32
ATGTCTTC	33
ATGTATGC	34

Selection of Cre mutants that recognize *lox* mutants on the inverted repeats region (halves) devoid of palindrome features was successfully performed. The mutations in such 5 *lox* mutants were in the flexibility region of *loxP*, namely, in positions 1 and/or 8-13 of each half. In some cases, *lox* mutants having mutations within the flexibility region of *loxP* were recognized by wt Cre.

In one embodiment, the present invention provides a composition of two Cre variants that can catalyze the recombination of non-palindromic sites. This compositions was also shown to catalyze asymmetric recombination also in sites which contain mutations within the restricted *intolerant* non-flexible sequence, i.e. in positions 2-7, thus demonstrating that 5 not only that the compositions and method of the invention are *not* restricted to the palindromic symmetry of recombination sites but they are also not to restricted to asymmetric recombination sites having mutations only within their flexible regions.

It was previously shown that the Cre protein sequence for DNA recognition of the inverted repeats *lox* site is independent of the protein sequence which is responsible of the 10 cleavage-ligation on the spacer *lox* site. Therefore, without wishing to be bound by any particular theory or mechanism of action, asymmetric recombination according to the present invention may be attributed to the fact that mutations within the docking recognition sites, namely the two halves that flank the spacer on the *lox* site, do not disrupt the cleavage-ligation mechanism. This attribution is supported by Shaikh et al. (*ibid*) who showed that 15 Cre/FLP chimeras successfully cleaved a chimeric substrate. The mode of cleavage of Cre was in *cis* whereas that of FLP was in *trans*. However, recombination of the chimeric protein did *not* occur.

In *vitro*, CM2 catalyzed the recombination of *loxP-M7* at a rather slower rate than the mixture of wt Cre and CM2. However, CM2 may not catalyze the recombination of *loxP-M7* 20 *in vivo*, as observed in the case of the *loxAT* mutant, which was recombined by wt Cre at a slower rate *in vitro*, but whose recombination was abolished *in vivo* (Hartung, et al., J Biol Chem 273:22884-22891, 1998). Expression of both wt Cre and CM2 catalyzed recombination of *loxP-M7*.

25 Example 2 - Selection strategy for Cre mutants that facilitate endogenous asymmetric recombination utilizing the *lox-LTR* recombination site

In retroviruses, *LTR* sequences flank the virus sequence from both sides. Lee et al. (Lee Y, et al., Biochem. Biophys. Res. Commun. 253:588-93, 1998) found a sequence 30 within the *LTR* which has some homology to *loxP*, this sequence is also termed "*lox-LTR*" (FIG. 4; SEQ ID NO:38). The spacer region of the *lox-LTR* was cloned into a *loxP* replacing the wt spacer and the new substrate was recognized by wt Cre to facilitate recombination.

Since *lox-LTR* sites are asymmetric, the selection strategy for Cre mutants that can recognize and catalyze specific recombination in these *lox-LTR* sites begins with the design of two *lox-LTR* derivatives as follows: In the first derivative, the half *left* site was placed on both sides of the spacer, in opposite directions (FIG. 5A) and in the second derivative, the 5 *right* half of the site was put on both sides of the spacer, in opposite directions (FIG. 5B). Each of these two symmetric *lox-LTR* derivatives was cloned within a construct in opposite directions on both sides of an antibiotic selectable gene (see FIGS. 5A-5B). In this construct, the *lox-LTR* site was inserted between a bacterial promoter and an antibiotic 10 resistance gene, which was positioned in opposite direction to the promoter (FIGS. 5A and 5B).

For the selection of new Cre mutants, the two constructs described in FIGS. 5A-5B are transformed into a Cre mutants library cells (Santoro et al., *ibid*). The antibiotic resistance gene is positioned in an opposite orientation to the promoter, and therefore is not transcribed. However, in the presence of competent Cre mutants, inversion recombination is 15 facilitated (FIG. 5B), giving rise to transcription and conferring antibiotic resistance to the bacteria. Positive Cre mutants that mediated recombination of constructs as described in FIGS. 5A and 5B are tested, individually and as a combination, in-vitro and in-vivo. The combination of a plurality of Cre mutants forms a heterotetrameric structure thereby facilitating recombination of a natural, asymmetric *lox-LTR* (FIG. 5C).

20

Example 3 - Selection of new Cre variants that facilitate integration of exogenous DNA into endogenous *lox*-like site.

Plant and animal genomes may contain *lox*-like sequences. Such sequences might be either symmetric or asymmetric. Due to dissimilarities between *lox*-like sequences and *loxP* 25 wt sequences, the *lox*-like sequences are not recognized by the wt Cre recombinase. Selection of Cre mutants that recognize endogenous *lox*-like sequences and facilitate site-specific integration of exogenous DNA into such sites requires initially in silico screening of databases containing the genome of desired species and to identify *lox*-like sequences. Screening is based on the homology of the *lox*-like spacer region (Lee et al, Gene, *ibid* and 30 Table 2).

After identifying endogenous *lox*-like sequences, a selection for Cre mutants that recognize the endogenous *lox*-like sequence is initiated by designing two *lox*-like

derivatives as follows: a. the half left site is placed on both sides of the spacer selected from any one of SEQ ID NOS:1-34, in opposite directions (FIG. 5A). b. the half right site is placed on both sides of the spacer, in opposite directions (see FIG. 5B). Each of the two new symmetric endogenous *lox*-like derivatives is cloned in opposite direction on both sides of an antibiotic selectable gene (FIGS. 5A-B). In this construct, the endogenous *lox*-like site on the left is inserted between a bacterial promoter and an antibiotic resistance gene, which is positioned in opposite direction to the promoter.

For the selection of new Cre mutants, the two constructs described above are transformed into a library of cells comprising and expressing Cre mutants such as the library disclosed in Santoro et al. (*ibid*). Preferably, a new library is formed for this purpose, by combining two Cre mutants' libraries (Santoro et al. *ibid*) and shuffling their elements using DNA shuffling methods known in the art (e.g. US Patent Nos. 6,326,204; 6,479,652 and 6,489,145).

Selection of suitable cells expressing the desired recombinase is attributed to the orientation of the promoter. Since the antibiotic resistance gene is orientated 'wrongly', i.e. in an opposite orientation to the promoter, it is not transcribed unless inversion recombination is facilitated, as described in FIG. 5B, by a competent Cre mutant. Thus, a desired Cre mutant gives rise to transcription of the antibiotic resistance gene thereby conferring antibiotic resistance to the cell comprising thereof. The desired Cre mutants are then tested in-vitro and in-vivo, individually or as a combination of a plurality of desired Cre mutants. It is suggested that asymmetric recombination is facilitated by a plurality of desired Cre mutants since a heterotetrameric structure of the plurality of desired Cre mutants is formed.

The foregoing description of the specific embodiments will so fully reveal the general nature of the invention that others can, by applying current knowledge, readily modify and/or adapt for various applications, such specific embodiments without undue experimentation and without departing from the generic concept, and, therefore, such adaptations and modifications should and are intended to be comprehended within the meaning and range of equivalents of the disclosed embodiments. It is to be understood that the phraseology or terminology employed herein is for the purpose of description and not of

limitation. The means, materials, and steps for carrying out various disclosed functions may take a variety of alternative forms without departing from the invention.

## **CLAIMS**

1. A genetically modified host cell transformed by an isolated exogenous DNA molecule wherein the isolated exogenous DNA molecule is integrated by asymmetric recombination into a predetermined locus of the genome of the cell.
- 5 2. The genetically modified host cell according to claim 1, wherein the isolated exogenous DNA molecule is integrated into sites of the genome chosen from the group consisting of: 3' UTRs, 5' UTRs, polyA sites and gene promoters.
3. The genetically modified host cell according to claim 1, wherein the cell is a eukaryotic cell.
- 10 4. The genetically modified host cell according to claim 3, wherein the cell is selected from the group consisting of: yeast, plant cell, embryonic stem cell, mesenchymal cell, and haematopoietic progenitor cell.
5. A transgenic organism comprising at least one genetically modified host cell, the genetically modified host cell being transformed by an isolated exogenous DNA molecule wherein the isolated exogenous DNA molecule is integrated by asymmetric recombination into a predetermined locus of the genome of the cell.
- 15 6. The transgenic organism according to claim 5, selected from the group consisting of: plant, yeast, or a vertebrate.
7. The genetically modified host cell according to claim 1, further transformed by at least one expression vector comprising a polynucleotide sequence encoding 20 at least one recombinase capable of catalyzing the asymmetric recombination.
8. The genetically modified host cell according to claim 7, wherein the at least one expression vector comprises a polynucleotide sequence encoding a plurality of recombinases capable of catalyzing the asymmetric recombination.
- 25 9. The genetically modified host cell according to claim 7, wherein the host cell is capable of expressing the at least one recombinase.
10. The genetically modified host cell according to claim 8, wherein the host cells is capable of expressing the plurality of recombinases.
11. The genetically modified host cell according to any one of claims 7 to 8, 30 wherein at least one recombinase is a mutant recombinase.

12. The genetically modified host cell according to any one of claims 7 to 8, wherein the recombinase is a Cre mutant capable of performing recombination between two recombination sites wherein at least one recombination site comprises a spacer sequence selected from the group consisting of: SEQ ID NOS. 1-34.
- 5
13. The genetically modified host cell according to any one of claims 7 to 8, wherein the expression vector is a plasmid or a virus.
14. The genetically modified host cell according to any one of claims 13, wherein the polynucleotide sequence encoding the recombinase is operably linked to expression control sequences.
- 10
15. The genetically modified host cell according to claim 13, wherein the polynucleotide sequence encoding the at least one recombinase further comprises a promoter.
16. The genetically modified host cell according to claim 15, wherein the promoter is derived from bacteria, yeast, insect, animal, plant or virus.
- 15
17. The genetically modified host cell according to claim 15, wherein the promoter is selected from the group consisting of: E. coli *lac* and *trp* operons, the *tac* promoter, the bacteriophage  $\lambda$  promoter, bacteriophage T7 and SP6 promoters,  $\beta$ -actin promoter, insulin promoter, human cytomegalovirus (CMV) promoter, HIV-LTR, RSV-LTR, SV40 promoter, baculoviral polyhedrin and p10 promoter.
- 20
18. The genetically modified host cell according to claim 15, wherein the promoter is an inducible promoter.
19. The genetically modified host cell according to claim 18, wherein the promoter is selected from the group consisting of: tetracycline, heat shock, steroid hormone, heavy metal, phorbol ester, adenovirus E1A element, interferon, and serum inducible promoters.
- 25
20. A method for genetically modifying cells by catalyzing a site-specific insertion of a selected DNA segment into at least one predetermined locus of the genome of the cells, comprising:
- 30

- (a) providing a population of cells having a first endogenous recombination site at a predetermined genomic locus;
- (b) introducing into the population of cells an isolated DNA comprising a selected DNA segment and further comprising a second recombination site;
- (c) introducing into the population of cells at least one vector comprising a polynucleotide sequence encoding at least one recombinase, the at least one recombinase being capable of catalyzing recombination between the first endogenous recombination site and the second recombination site;
- (d) obtaining site-specific insertion of the selected DNA segment at the predetermined genomic locus; and
- (e) selecting a subpopulation of cells or a cell progeny from said population of cells, the subpopulation of cells is devoid of said at least one recombinase;

wherein said first endogenous recombination site is an asymmetric recombination site.

21. The method according to claim 20, wherein the first endogenous recombination site and the second recombination site are the same asymmetric recombination sites.
- 20 22. The method according to claim 20, comprising introducing into the population of cells at least one vector comprising a polynucleotide sequence encoding at least two distinct recombinases capable of catalyzing the asymmetric recombination.
23. The method according to claim 20, wherein the polynucleotide sequence encoding the at least one recombinase is operably linked to expression control sequences.
- 25 24. The method according to claim 20, the at least one vector further comprising a regulatory nucleotide sequence such that expression of said at least one recombinase is produced by activating the regulatory nucleotide sequence.
25. The method according to claim 20, wherein the polynucleotide sequence encoding the at least one recombinase further comprises a promoter.

26. The method according to claim 25, wherein the promoter is derived from bacteria, yeast, insect, animal, plant or virus.
27. The method according to claim 25, wherein the promoter is selected from the group consisting of: *E. coli lac* and *trp* operons, the *tac* promoter, the bacteriophage  $\lambda$  promoter, bacteriophage T7 and SP6 promoters,  $\beta$ -actin promoter, insulin promoter, human cytomegalovirus (CMV) promoter, HIV-LTR, RSV-LTR, SV40 promoter, baculoviral polyhedrin and p10 promoter.
- 5 28. The method according to claim 25, wherein the promoter is an inducible promoter.
- 10 29. The method according to claim 26, wherein the promoter is selected from the group consisting of: tetracycline, heat shock, steroid hormone, heavy metal, phorbol ester, adenovirus E1A element, interferon, and serum inducible promoters.
- 15 30. The method according to claim 20, wherein the selected DNA segment is a gene for a structural protein, an enzyme, or a regulatory molecule.
31. The method according to claim 20, the at least one vector is a plasmid or a virus.
32. A method for genetically modifying cells by catalyzing a site-specific excision of an endogenous DNA segment from a predetermined locus of the genome of the cells, comprising:
- 20 (a) providing a population of cells comprising an endogenous DNA segment within a predetermined locus of the genome thereof, the selected endogenous DNA segment is flanked by two recombination sites having the same orientation, wherein at least one recombination site is asymmetric;
- 25 (b) introducing into the population of cells a DNA molecule comprising a polynucleotide sequence encoding at least one recombinase, the at least one recombinase is capable of catalyzing recombination between the two recombination sites;
- (c) obtaining site-specific excision of the endogenous DNA segment from the predetermined locus of the genome of said population of cells; and

- (d) selecting a subpopulation of cells or a cell progeny from said population of cells, the subpopulation of cells is devoid of said at least one recombinase.
33. The method according to claim 32, wherein the DNA molecule further comprising a regulatory nucleotide sequence and expression of the at least one recombinase is produced by activating the regulatory nucleotide sequence.
- 5 34. The method according to claim 32, wherein the selected endogenous DNA segment is a gene for a structural protein, an enzyme, or a regulatory.
35. The method according to claim 32, the two recombination sites have opposite orientations and the site-specific recombination results in a site-specific inversion of the nucleotide sequence of the selected endogenous DNA segment.
- 10 36. The method according to claim 32, wherein the polynucleotide sequence encoding the at least one recombinase further comprises a promoter.
37. The genetically modified host cell according to claim 36, wherein the promoter is derived from bacteria, yeast, insect, animal, plant or virus.
- 15 38. The method according to claim 36, wherein the promoter is selected from the group consisting of: *E. coli lac* and *trp* operons, the *tac* promoter, the bacteriophage  $\lambda$  promoter, bacteriophage T7 and SP6 promoters,  $\beta$ -actin promoter, insulin promoter, human cytomegalovirus (CMV) promoter, HIV-LTR, RSV-LTR, SV40 promoter, baculoviral polyhedrin and p10 promoter.
- 20 39. The method according to claim 36, wherein the promoter is an inducible promoter.
40. The method according to claim 39, wherein the promoter is selected from the group consisting of: tetracycline, heat shock, steroid hormone, heavy metal, phorbol ester, adenovirus E1A element, interferon, and serum inducible promoters.
- 25 41. A method for catalyzing asymmetric recombination, comprising:
- 30 (a) providing a first composition comprising at least one DNA molecule, the at least one DNA molecule comprising two recombination sites, wherein at least one recombination site is an asymmetric recombination site;



## **ABSTRACT**

The present invention relates to compositions and methods for catalyzing recombination of non-palindromic recombination sites in a cell free system, in isolated cells or in living organisms. The compositions and methods of the invention are suitable for mediating specific or random recombinations between DNA sequences comprising at least one non-palindromic recombination site.

10

ATAACTTCGTATAGCATACATTATACGAAGTTAT WT *loxP*  
 SEQ ID NO:35

ATAACTCTATAGCATACATTATAGAGTTAT WT *loxM7*  
 SEQ ID NO:36

ATAACTTCGTATAGCATACATTATAGAGTTAT *loxP-M7*  
 SEQ ID NO:37

Figure 1A

*loxP*

<b>WT <i>loxP</i></b> <b>SEQ ID NO:39</b>	<b>XhoI</b> Xho-Bam [CTCGAG]ATAACT <u>TCGTATAGCATACATTATACGAAGTTAT</u> GGATTC Bam-Xho GAGCTCTATTGAAGCATATCGTATGTAATATGCTTCATAACCTAGG	<b>BamHI</b> GGATTC
	<b>EcoRI</b> Eco-Pst [GAATTCA]ATAACT <u>TCGTATAGCATACATTATACGAAGTTAT</u> CTGCAG Pst-Eco CTTAAGTATTGAAGCATATCGTATGTAATATGCTTCATAAGACGTC	<b>PstI</b> CTGCAG

*WT *loxM7**  
**SEQ ID NO:40**

<b>XhoI</b> Xho-Bam [CTCGAG]ATAACT <u>CTATAGCATACATTATAGAGTTAT</u> GGATTC Bam-Xho GAGCTCTATTGAGATATCGTATGTAATATCTCAATAACCTAGG	<b>BamHI</b> GGATTC	
	<b>EcoRI</b> Eco-Pst [GAATTCA]ATAACT <u>CTATAGCATACATTATAGAGTTAT</u> CTGCAG Pst-Eco CTTAAGTATTGAGATATCGTATGTAATATCTCAATAAGACGTC	<b>PstI</b> CTGCAG

*WT *loxP-M7**  
**SEQ ID NO:41**

<b>XhoI</b> Xho-Bam [CTCGAG]ATAACT <u>TCGTATAGCATACATTATAGAGTTAT</u> GGATTC Bam-Xho GAGCTCTATTGAAGCATATCGTATGTAATATCTCAATAACCTAGG	<b>BamHI</b> GGATTC	
	<b>EcoRI</b> Eco-Pst [GAATTCA]ATAACT <u>TCGTATAGCATACATTATAGAGTTAT</u> CTGCAG Pst-Eco CTTAAGTATTGAAGCATATCGTATGTAATATCTCAATAAGACGTC	<b>PstI</b> CTGCAG

Figure 1B

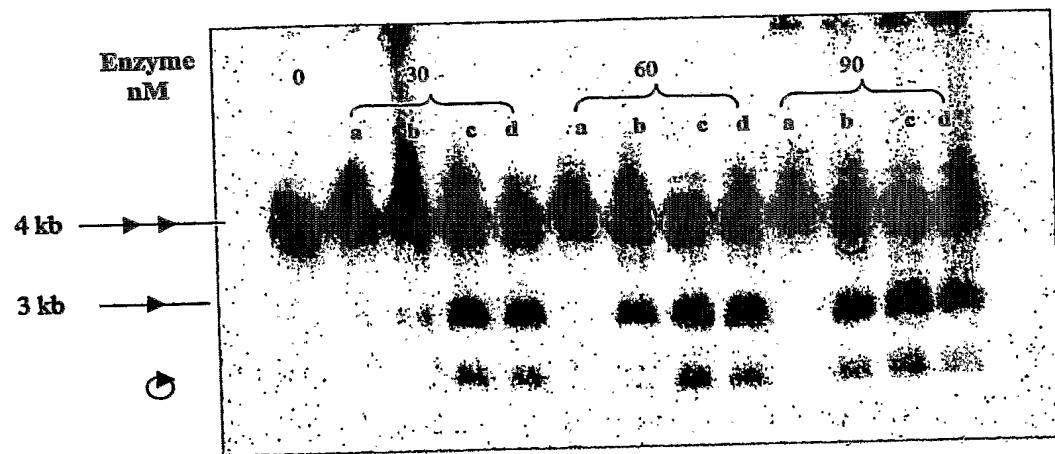


Figure 2A

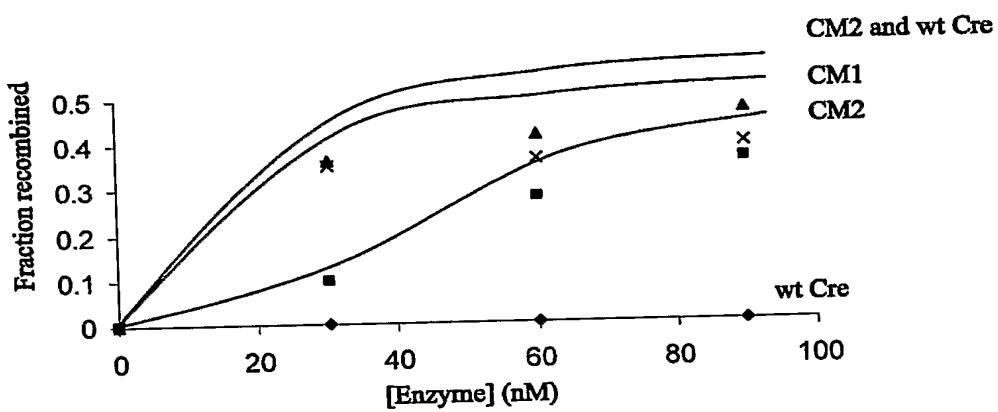


Figure 2B

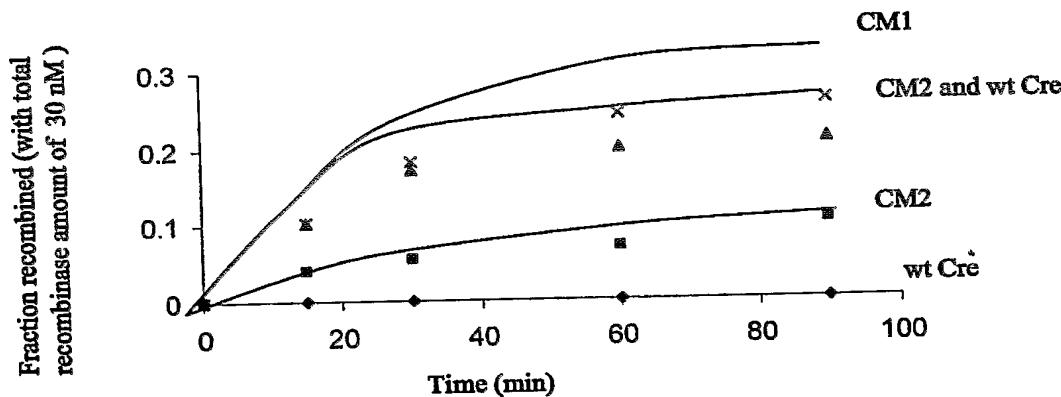


Figure 3A

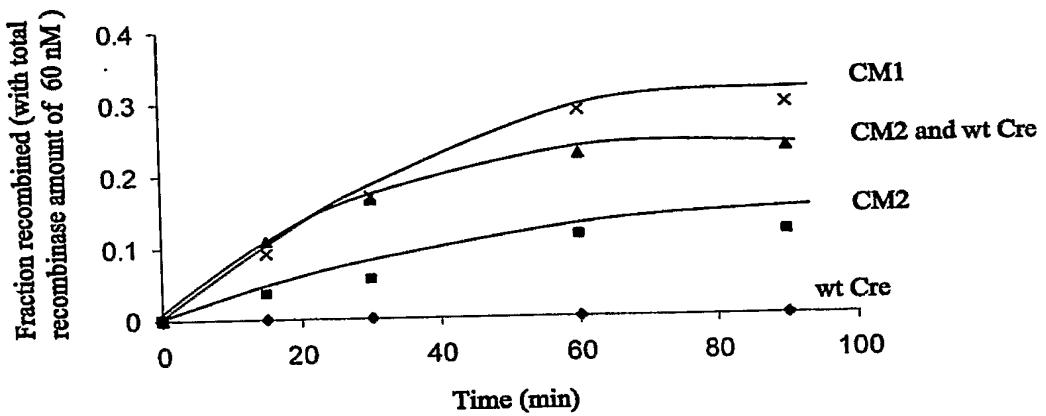


Figure 3B

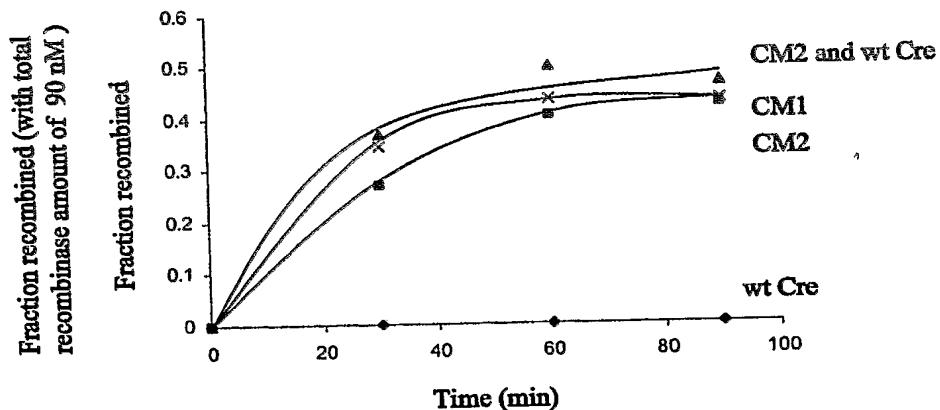


Figure 3C

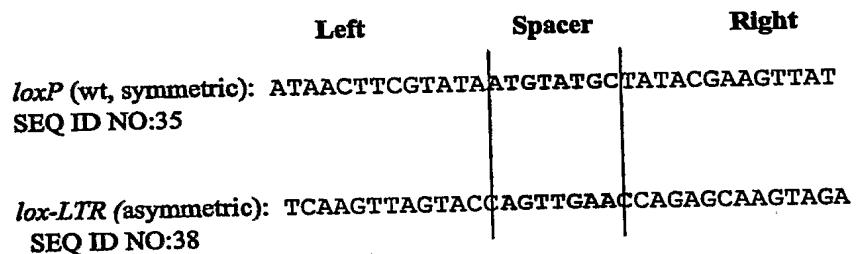


Figure 4

KIDUM/005 US Prov.

5/7

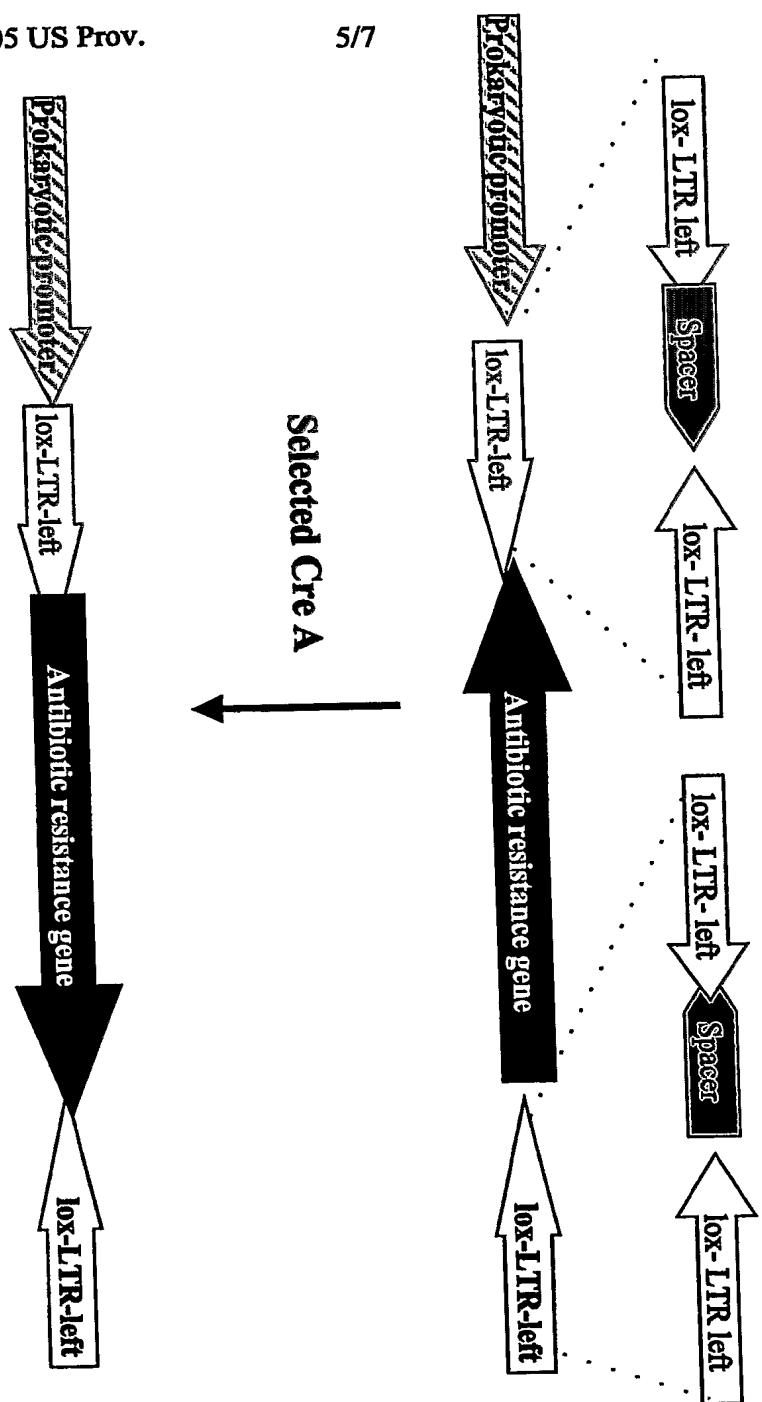
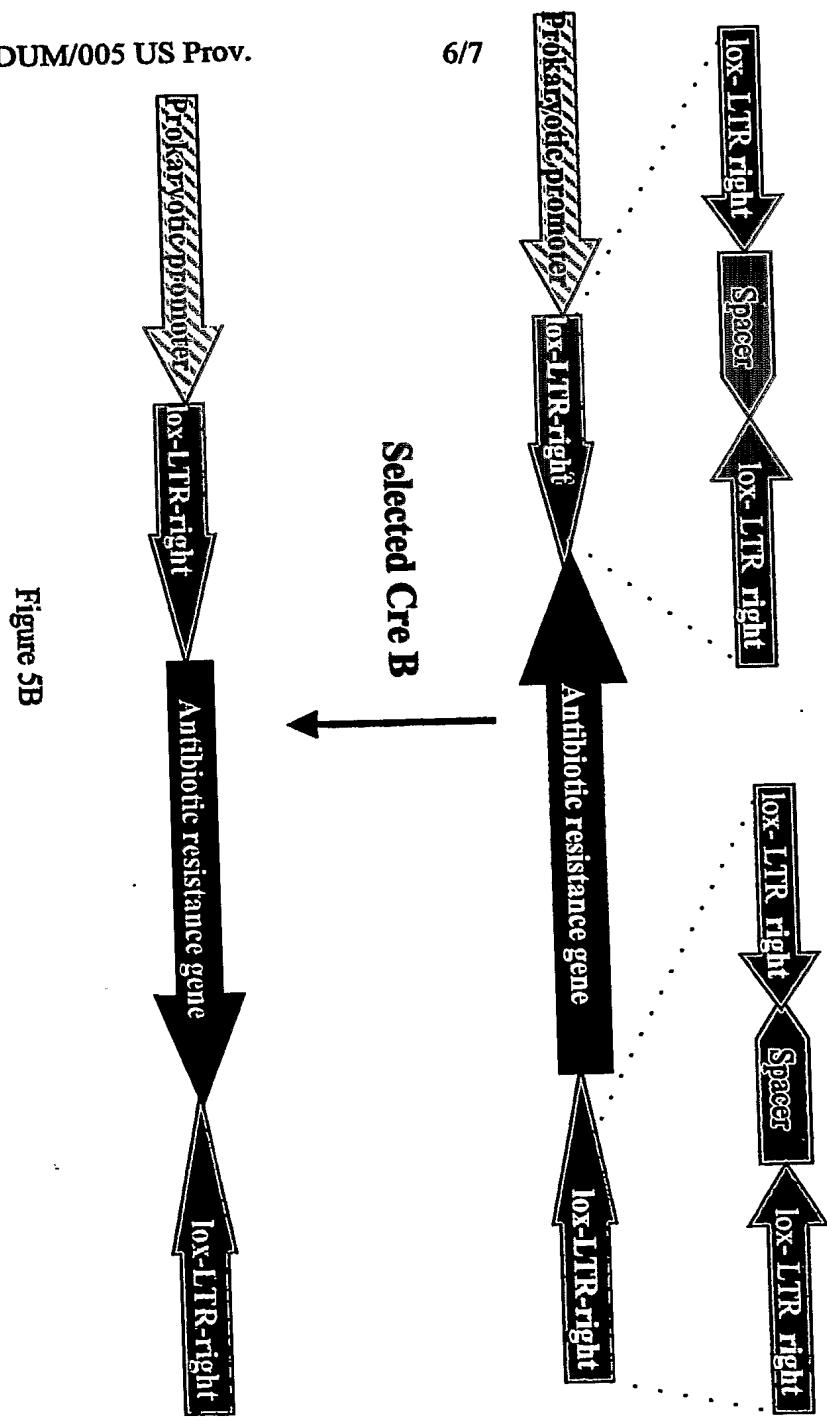


Figure 5A

KIDUM/005 US Prov.

6/7



KIDUM/005 US Prov.

7/7

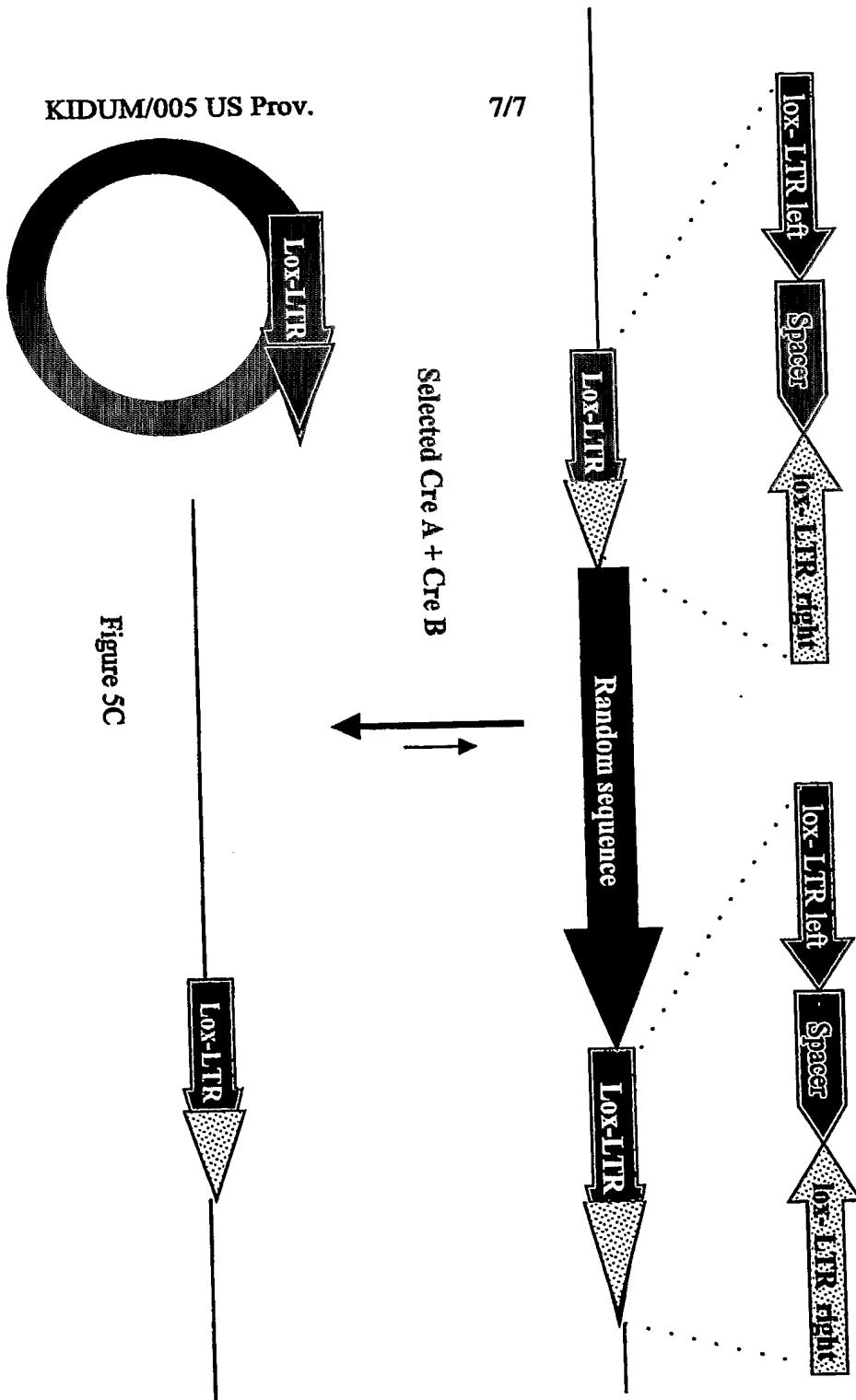


Figure 5C